MLE for Categorical

Data D: N samples, K categories, N_k samples in the k-th.

like.

$$L(\theta; D) \propto \prod_{i \in [K]} \theta_i^{N_i} \quad \text{where} \quad \sum \theta_i = 1.$$

log-like

$$\ell(\theta; D) = \sum N_i \log \theta_i + C$$

Let (Lagrange mult.)

$$f(\theta) = \sum N_i \log \theta_i + \lambda(1 - \sum \theta_i)$$

$$\frac{\partial f}{\partial \lambda} = 1 - \sum \theta_i \text{ and } \frac{\partial f}{\partial \theta_i} = \frac{N_i}{\theta_i} - \lambda$$

all set to $0 \Rightarrow \sum \theta_i = 1$

$$\frac{\nu_i}{\theta_i} = \lambda \text{ and } \lambda \nu_i = \lambda \theta_i \quad \forall i$$

So $\lambda = \lambda \sum \theta_i = \sum \lambda \theta_i = \sum \nu_i = N$

and $\theta_i = \frac{\nu_i}{\lambda} = \frac{\nu_i}{N}$